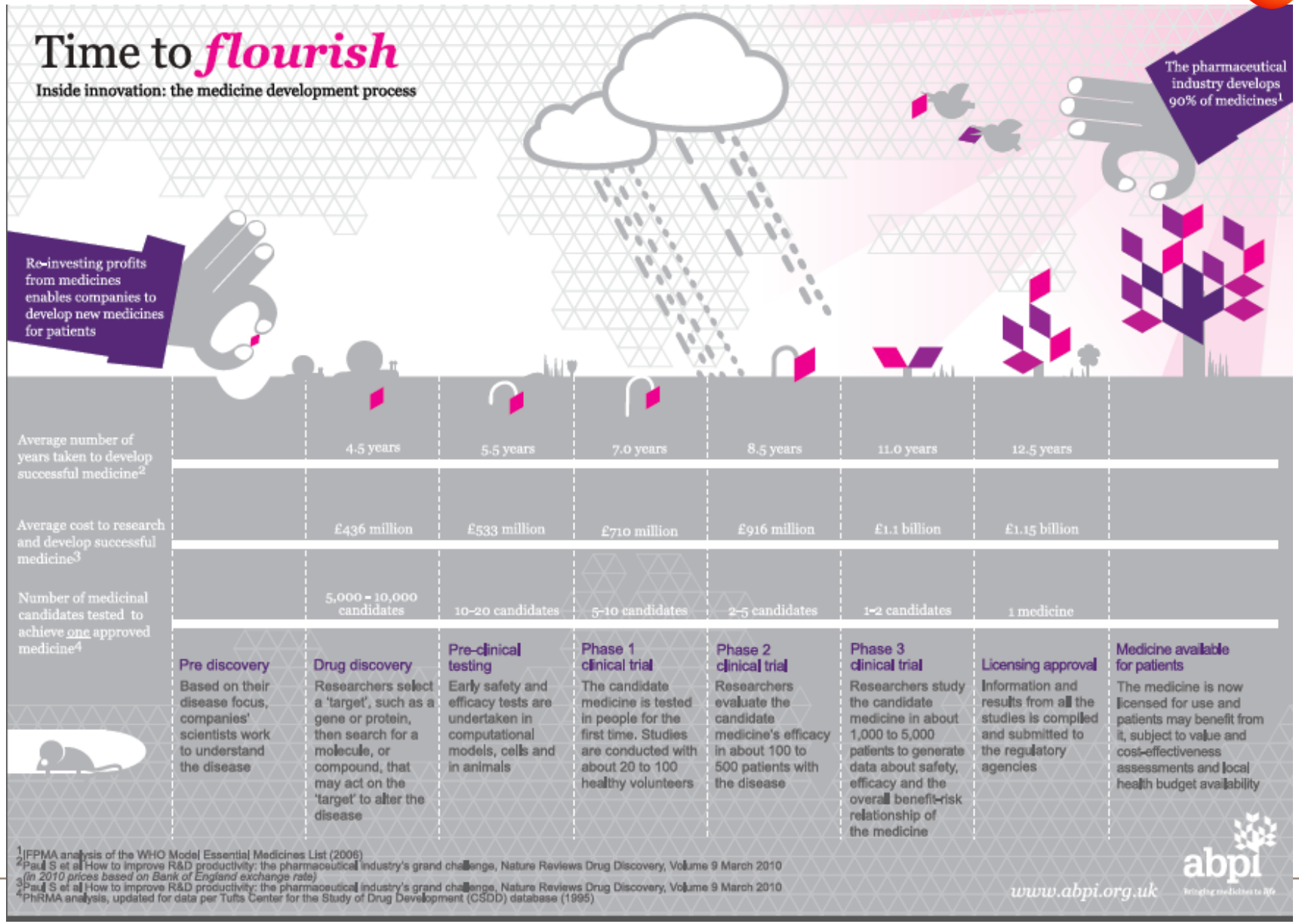
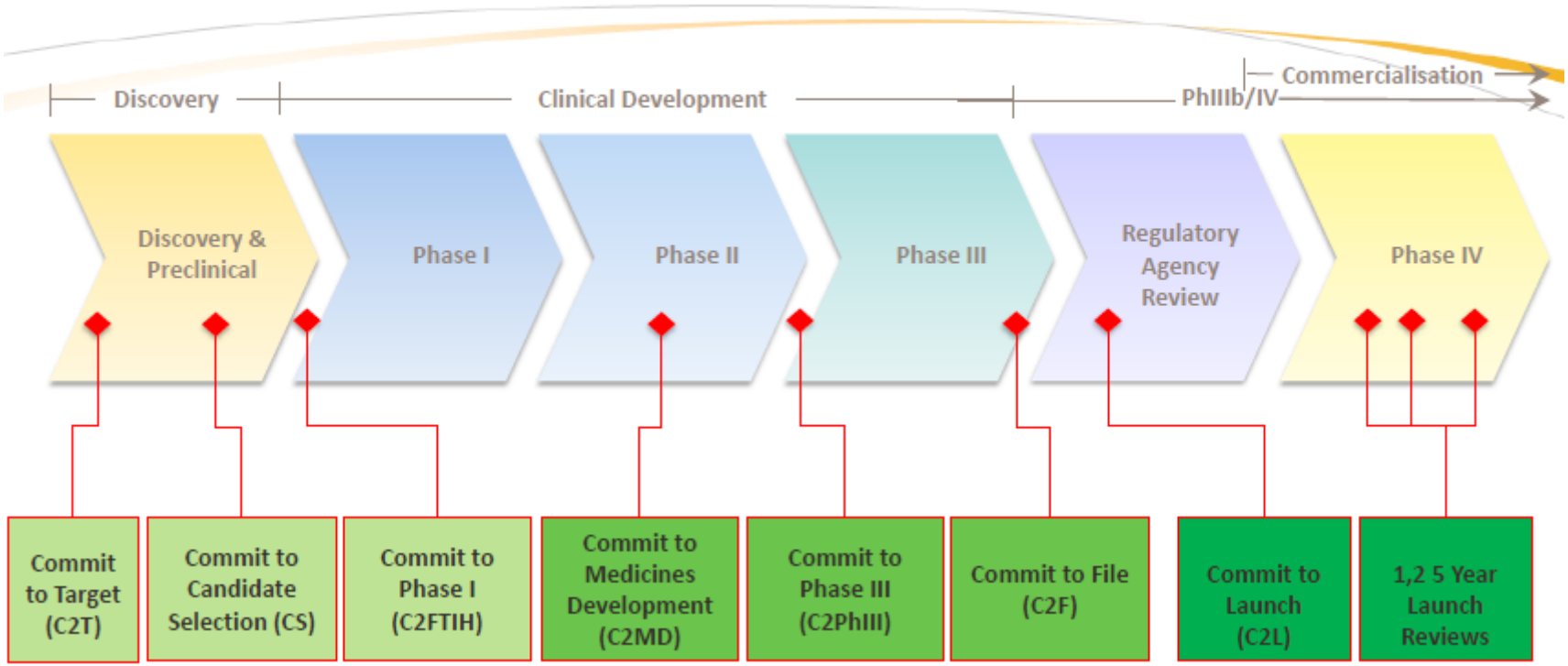
Using prior elicitation and Bayesian thinking to help shape decision making in the pharmaceutical industry

Nicky Best
Statistical Innovation Group, GSK

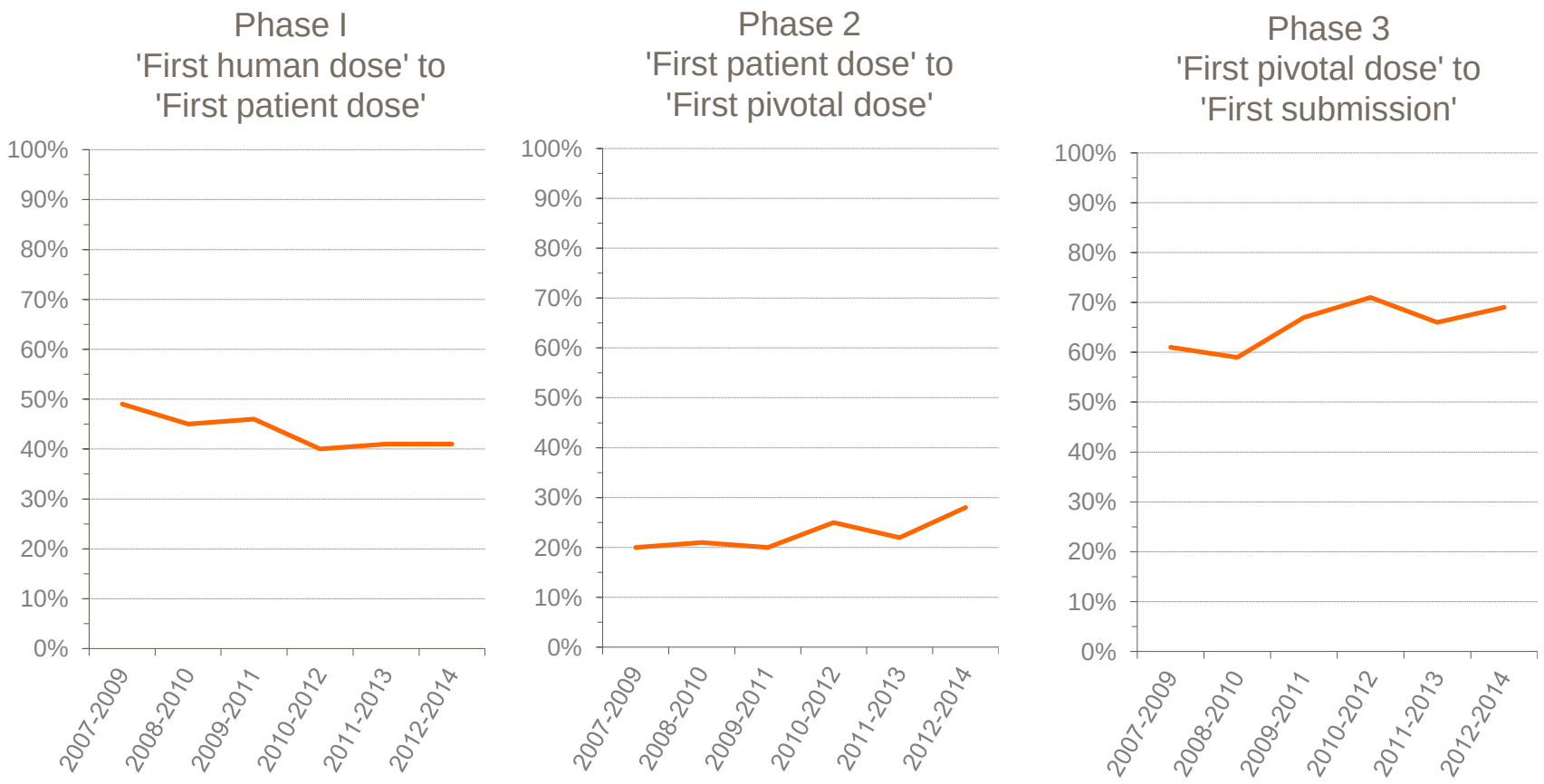
The Drug Development Process



Key Milestone Decisions Gates Through Drug Development



Trends in Pharmaceutical Industry Success Rates



Based on data from a consistent cohort of 20 companies participating each year between 2008 and 2015.

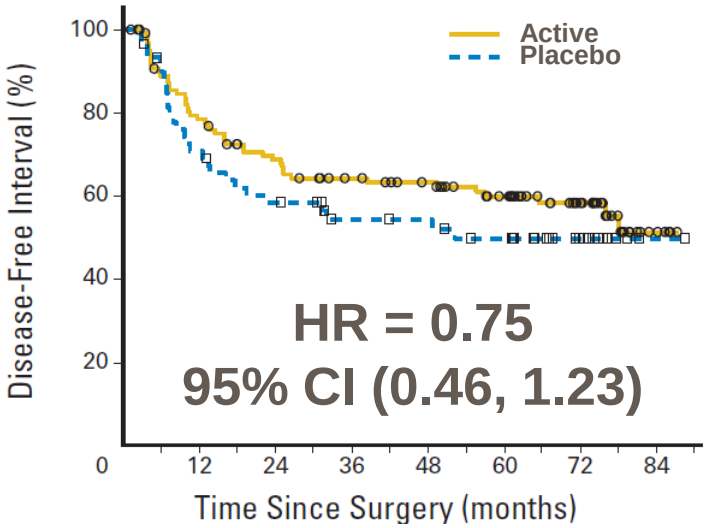
The power and the glory.....or yet another failure

Most late phase clinical trials are conducted with 90% power, but the success rate is much less than 90%

Why is this?

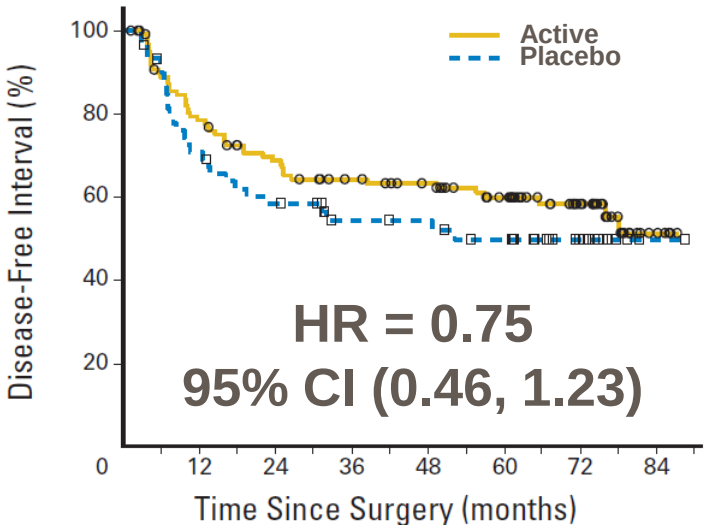
Case study: Cancer trial

Phase 2 study results



Case study: Cancer trial

Phase 2 study results

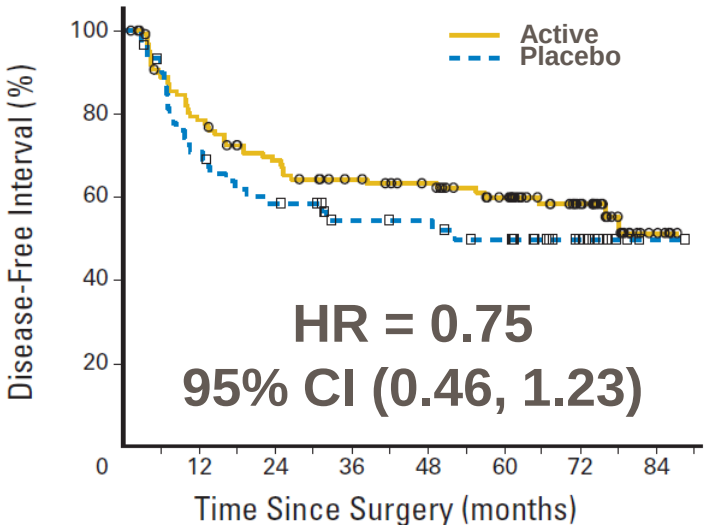


Phase 3 study results

HR = 1.02
95% CI (0.89, 1.18)

Case study: Cancer trial

Phase 2 study results



How can we better discharge risk?

Phase 3 study results

HR = 1.02
95% CI (0.89, 1.18)

Should we be surprised?

Power is not knowledge

A protocol might say something like this ...

Assuming a clinically relevant difference of 2 points on the primary endpoint scale, with a standard deviation of 6.2, 200 subjects per arm are required to provide 90% power at the 5% alpha level (two-sided).

Power is not knowledge

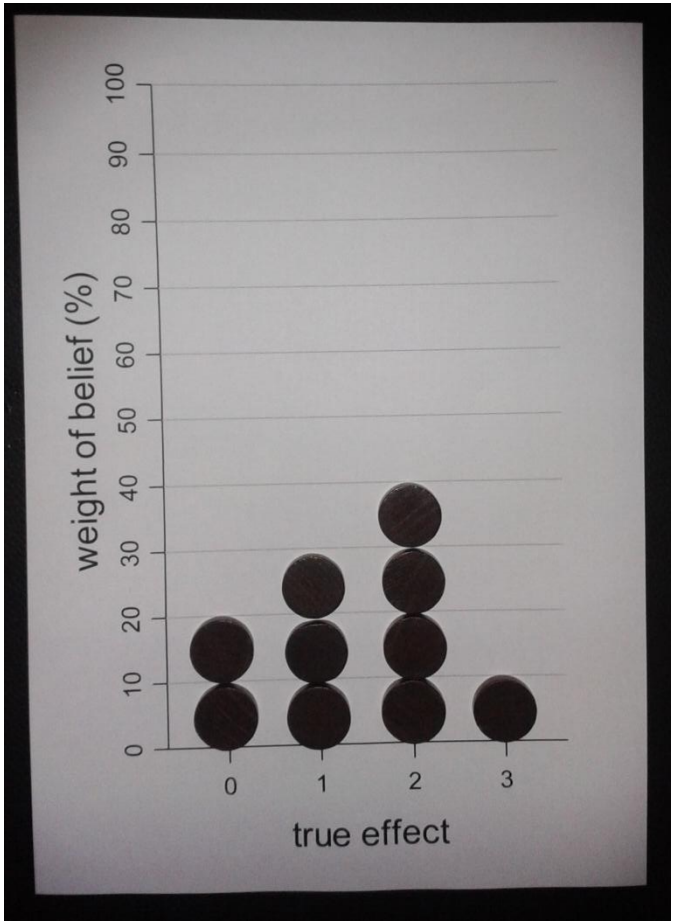
A protocol might say something like this ...

Assuming a clinically relevant difference of 2 points on the primary endpoint scale, with a standard deviation of 6.2, 200 subjects per arm are required to provide 90% power at the 5% alpha level (two-sided).

We are assuming with 100% certainty that the true effect of the drug is 2 points.

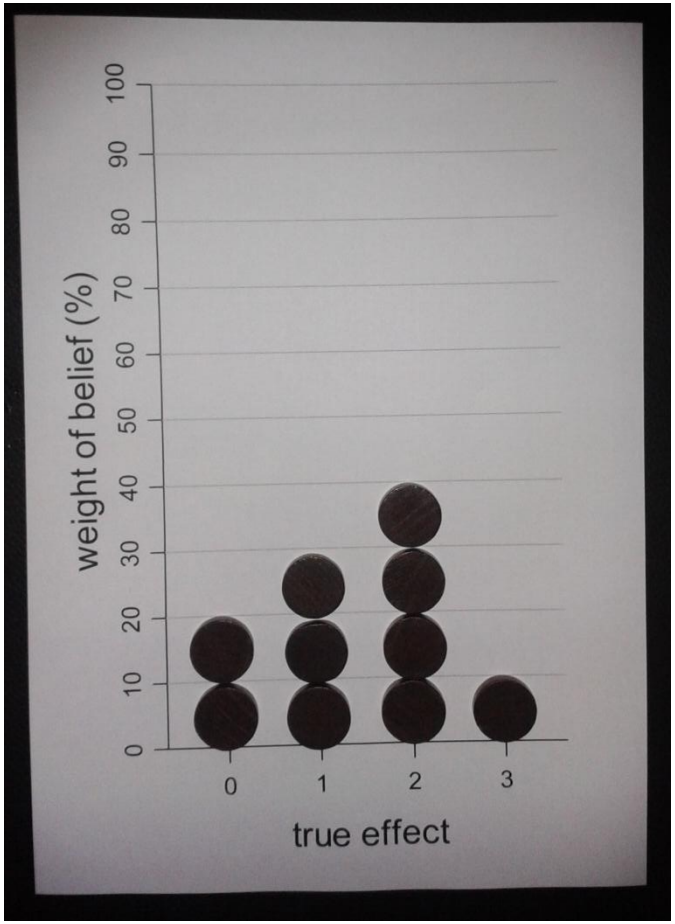
Power is not knowledge

Expert belief about true effect

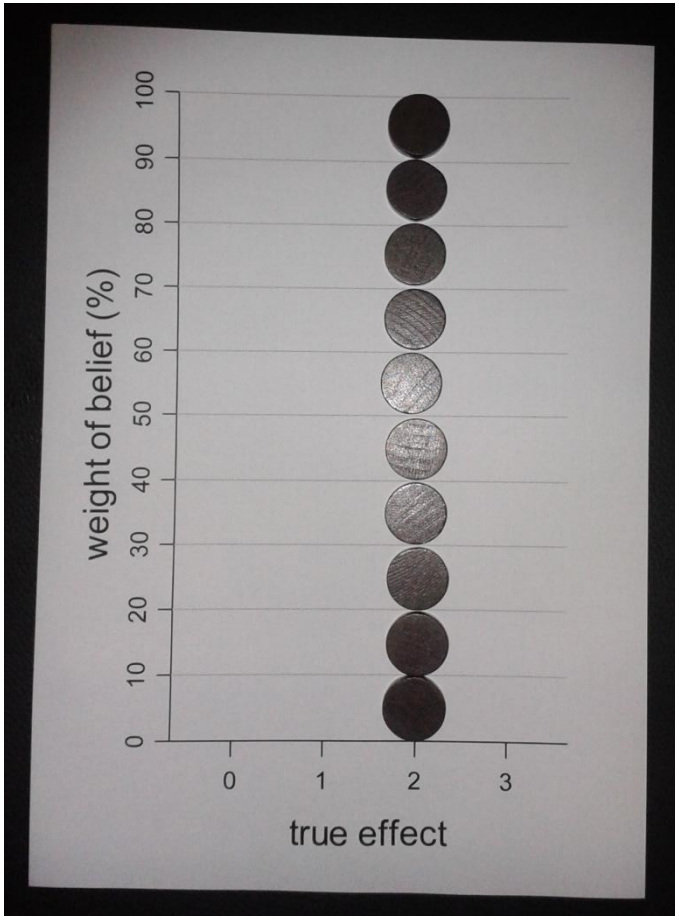


Power is not knowledge

Expert belief about true effect



Power calculation assumption



Power and Assurance

True effect size	Power	Expert Belief
0	2.5%	20%
1	36%	30%
2	90%	40%
3	99.8%	10%

Power and Assurance

True effect size	Power	Expert Belief
0	2.5%	20%
1	36%	30%
2	90%	40%
3	99.8%	10%

Power = 90%

the probability of success assuming the true (unknown and never known) effect of the drug is 2 points

Power and Assurance

True effect size	Power	Expert Belief	Power x Belief	
0	2.5%	20%	0.5%	
1	36%	30%	10.8%	**57%**
2	90%	40%	36%	
3	99.8%	10%	9.9%	

Power = 90%

the probability of success assuming the true (unknown and never known) effect of the drug is 2 points

Power and Assurance

True effect size	Power	Expert Belief	Power x Belief
0	2.5%	20%	0.5%
1	36%	30%	10.8%
2	90%	40%	36%
3	99.8%	10%	9.9%

57%

Power = 90%

the probability of success assuming the true (unknown and never known) effect of the drug is 2 points

BUT

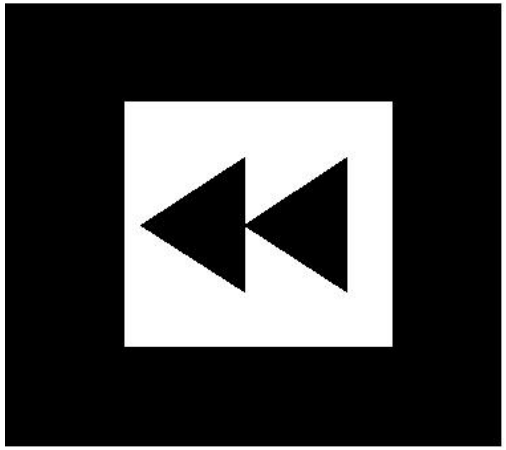
Assurance (prob of success) = 57%

the average of the power calculations, weighted by the belief about how big the true effect size is

Back to the cancer trial....Lets travel back in time!

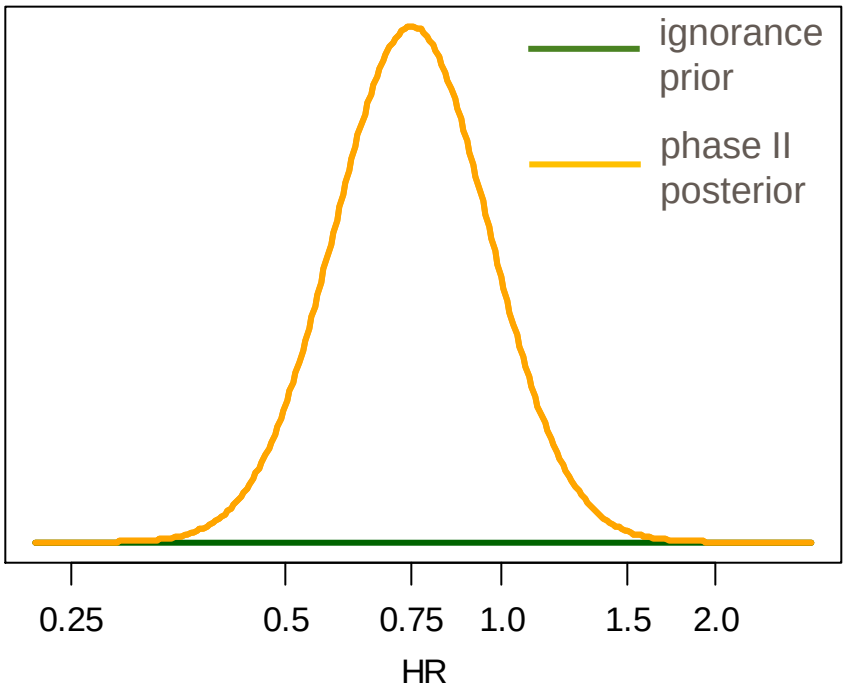
– What would you like to know before doing the study that would help you make an investment decision?

– Rewind 10 years



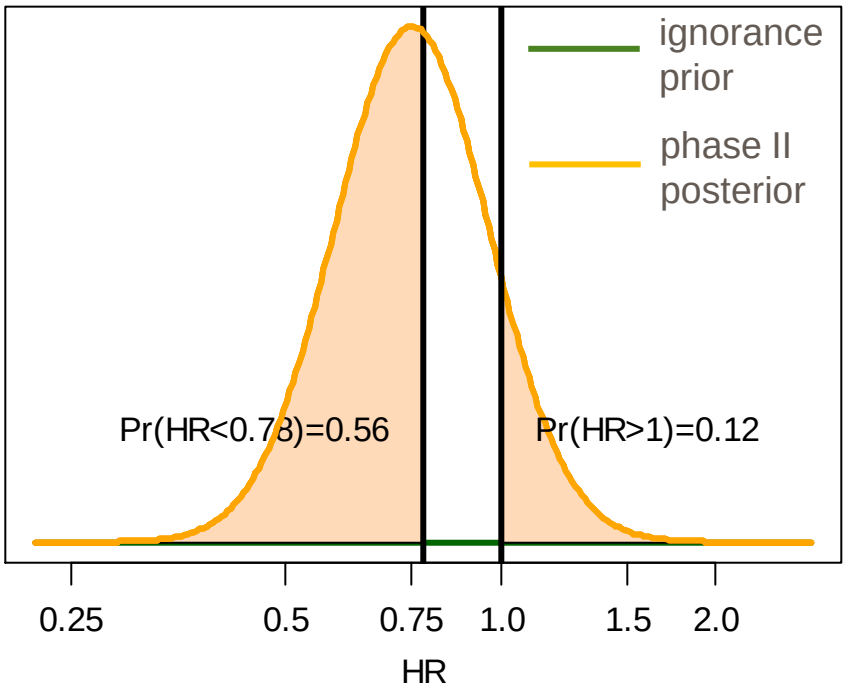
Back to the cancer trial....

- Designed to have 90% power to detect clinically relevant HR of 0.78
- What do the Phase II data tell us about the treatment effect?
  - Conventional frequentist analysis gives HR = 0.75; 95% CI (0.46, 1.23)
  - Bayesian analysis with ‘ignorance’ prior:



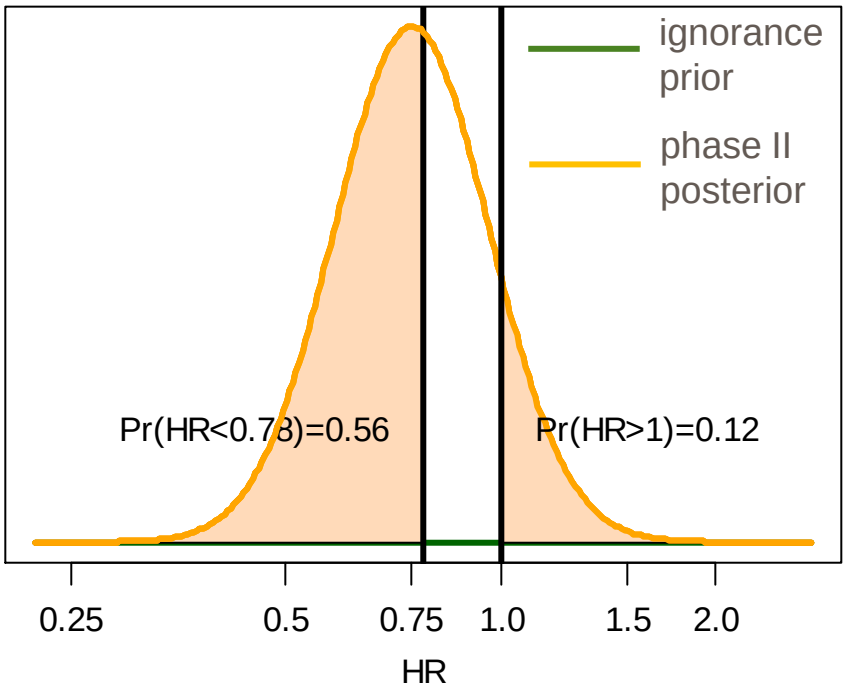
Back to the cancer trial....

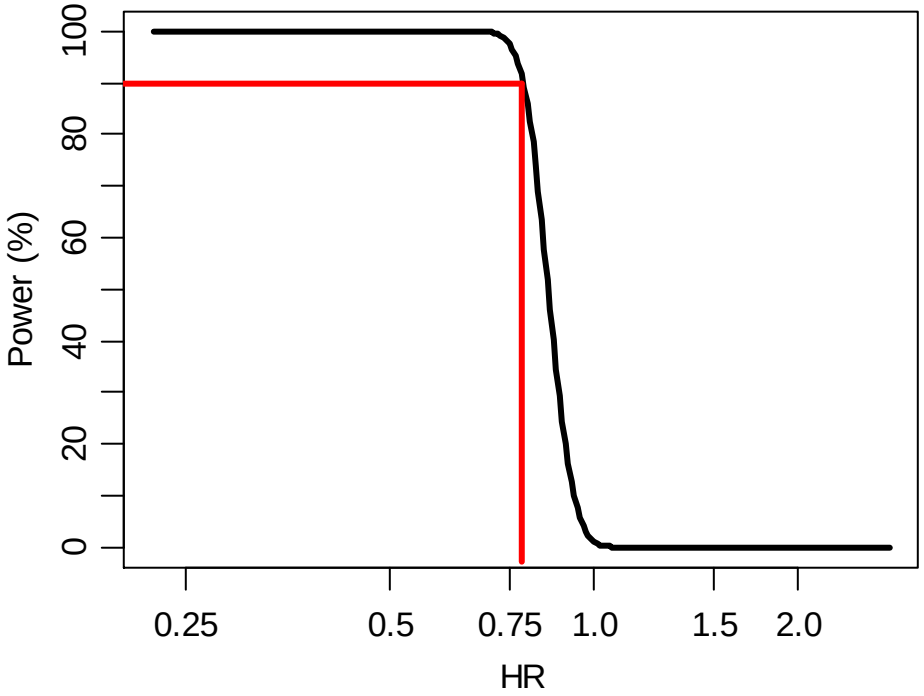
- Designed to have 90% power to detect clinically relevant HR of 0.78
- What do the Phase II data tell us about the treatment effect?
  - Conventional frequentist analysis gives HR = 0.75; 95% CI (0.46, 1.23)
  - Bayesian analysis with 'ignorance' prior:



Back to the cancer trial....

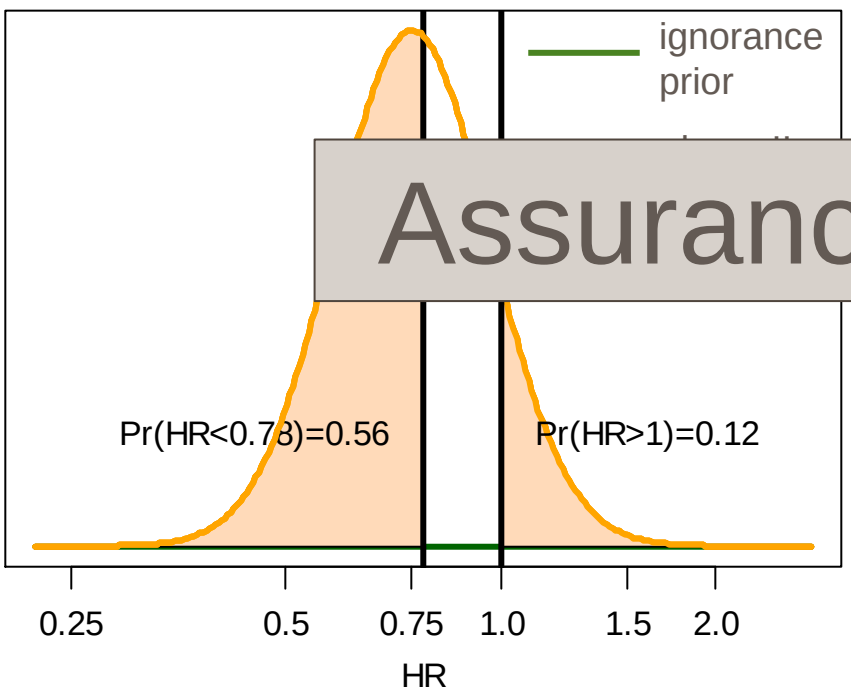
- Designed to have 90% power to detect clinically relevant HR of 0.78
- What do the Phase II data tell us about the treatment effect?
  - Conventional frequentist analysis gives HR = 0.75; 95% CI (0.46, 1.23)
  - Bayesian analysis with 'ignorance' prior:



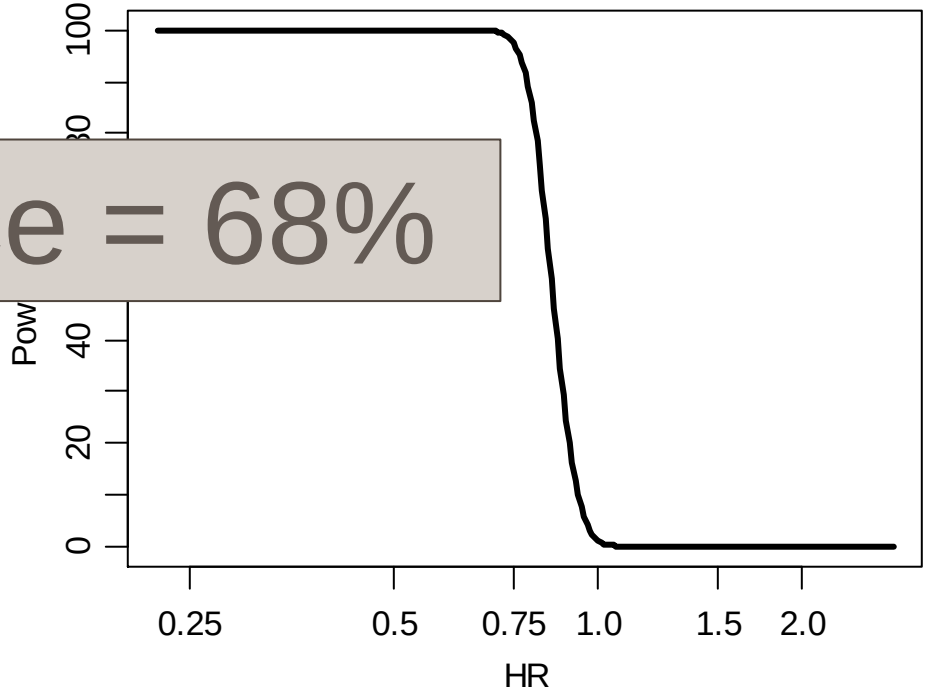


Back to the cancer trial....

- Designed to have 90% power to detect clinically relevant HR of 0.78
- What do the Phase II data tell us about the treatment effect?
  - Conventional frequentist analysis gives HR = 0.75; 95% CI (0.46, 1.23)
  - Bayesian analysis with 'ignorance' prior:



Assurance = 68%



Back to the cancer trial....

Assurance = 68%

- Probability that the trial will meet its primary endpoint based on current (....we are still back in time...) evidence about the treatment effect

- Is this probability high or low?
- Phase 2 trial does not exist in a vacuum – what other evidence should we take into account to produce our prior?
- Phase 3 setting ≠ Phase 2 setting
  - Different treatments
  - Different populations

Uncertainty is not Ignorance

- Even if we have only imperfect knowledge about an asset
  - How it performed in a related population
  - What our competitors have found with the same mechanism
  - What I know about the disease (which you might not know)

  → this can be used to help interrogate potential future clinical designs

Uncertainty is not Ignorance

- Even if we have only imperfect knowledge about an asset
  - How it performed in a related population
  - What our competitors have found with the same mechanism
  - What I know about the disease (which you might not know)

  → this can be used to help interrogate potential future clinical designs
- We do this by formally combining knowledge and data, into a "prior distribution", that represents our best expression of what is known, "just now", about the true drug effect of our asset

Uncertainty is not Ignorance

- Even if we have only imperfect knowledge about an asset
  - How it performed in a related population
  - What our competitors have found with the same mechanism
  - What I know about the disease (which you might not know)

  → this can be used to help interrogate potential future clinical designs
- We do this by formally combining knowledge and data, into a "prior distribution", that represents our best expression of what is known, "just now", about the true drug effect of our asset
- The prior can be used to interrogate potential clinical trial designs and development plans, in order to assess their utility
  - Which of three trial designs has the highest probability of success?
  - Should we incorporate an interim futility test, because our current state of knowledge is too diffuse?
  - Should we go straight to Phase 3? Do we believe enough in our drug now to make that commitment?

GSK Prior Elicitation Initiative

- Prior knowledge exists on every project in some form
  - Different levels of uncertainty in predictability or relevance of the information
  - Often a translational gap between historical and current settings

GSK Prior Elicitation Initiative

- Prior knowledge exists on every project in some form
  - Different levels of uncertainty in predictability or relevance of the information
  - Often a translational gap between historical and current settings
- GSK have implemented formal expert elicitation to translate this information into quantitative prior distributions

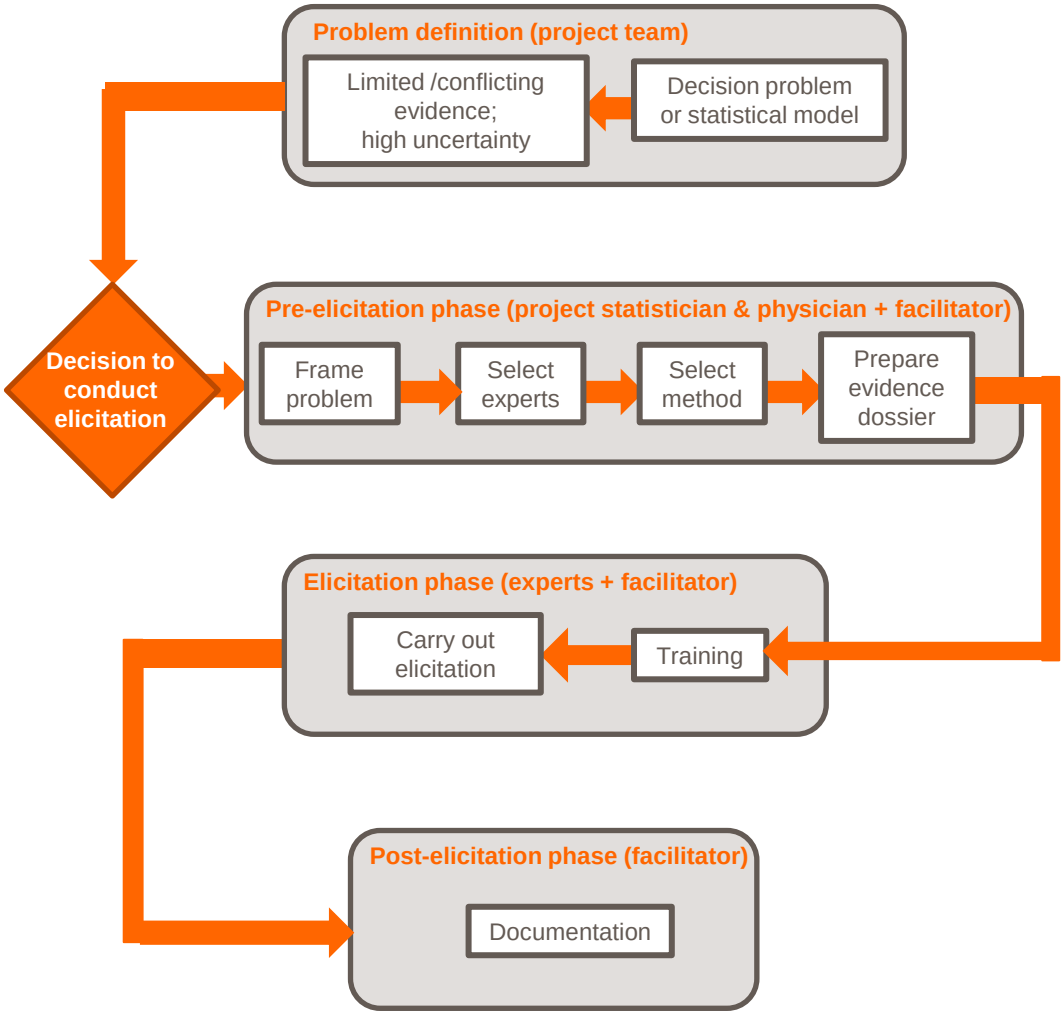
GSK Prior Elicitation Initiative

- Prior knowledge exists on every project in some form
  - Different levels of uncertainty in predictability or relevance of the information
  - Often a translational gap between historical and current settings
- GSK have implemented formal expert elicitation to translate this information into quantitative prior distributions
- Elicited priors used to:
  - determine assurance (prior predictive probability of success for a future study)
  - design clinical trials (e.g. plan interims, compare development strategies, stagger investment)
  - draw statistical inference (i.e. analysis of study data)

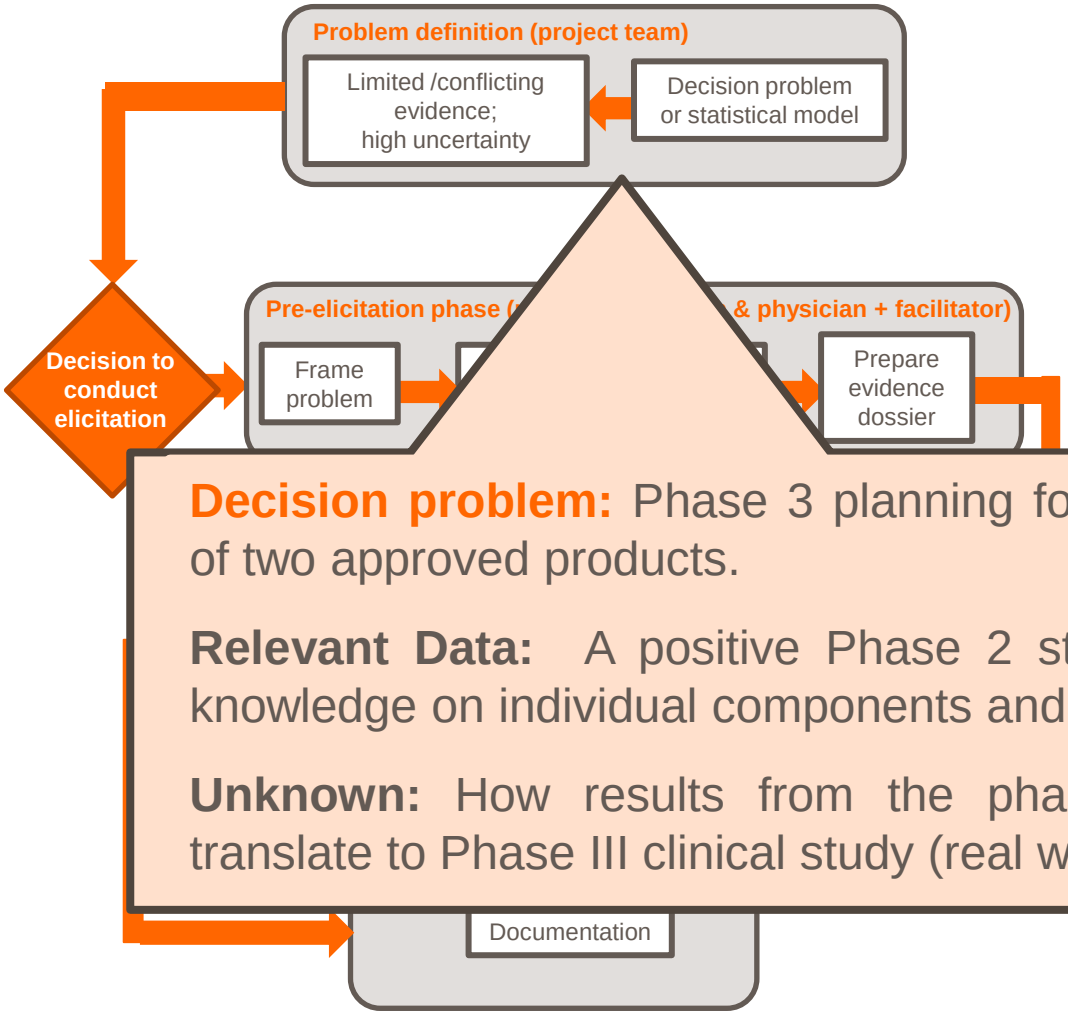
GSK Prior Elicitation Initiative

- Prior knowledge exists on every project in some form
  - Different levels of uncertainty in predictability or relevance of the information
  - Often a translational gap between historical and current settings
- GSK have implemented formal expert elicitation to translate this information into quantitative prior distributions
- Elicited priors used to:
  - determine assurance (prior predictive probability of success for a future study)
  - design clinical trials (e.g. plan interims, compare development strategies, stagger investment)
  - draw statistical inference (i.e. analysis of study data)
- Additional by-products of the elicitation process include:
  - Dedicated time for team to discuss all relevant data
  - Transparency of beliefs and rationale for those beliefs
- Enables uncertainty to be appropriately be captured and communicated

GSK Prior Elicitation process



Example of Prior Elicitation at GSK



Decision problem: Phase 3 planning for fixed dose combination (FDC) of two approved products.

Relevant Data: A positive Phase 2 study and a wealth of data and knowledge on individual components and other FDCs.

Unknown: How results from the phase 2 study (challenge model) translate to Phase III clinical study (real world situation).

Example of Prior Elicitation at GSK

Problem definition (project team)

Limited /conflicting evidence; high uncertainty

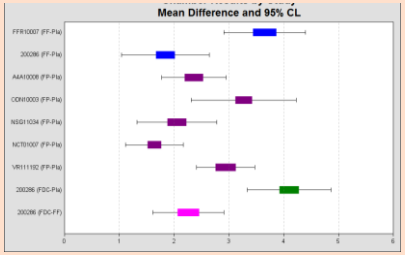
Decision problem or statistical model

Elicitation aim: to elicit true mean treatment difference between FDC and monotherapy

Decision to conduct elicitation

Pre-elicitation phase (project statistician & physician + facilitator)

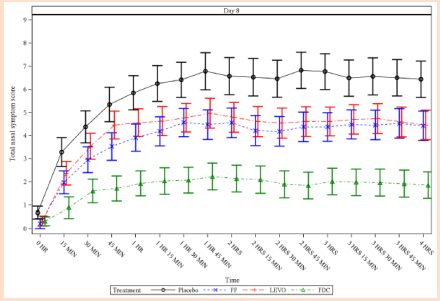
Frame problem → Select experts → Select method → Prepare evidence dossier



Data summaries from GSK reports and published competitor studies

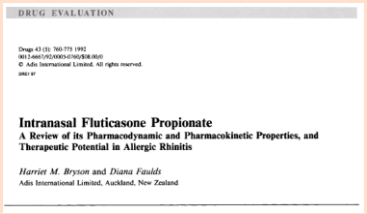
Evidence dossier

Regulatory Reviews



GSK Historical Data Sets

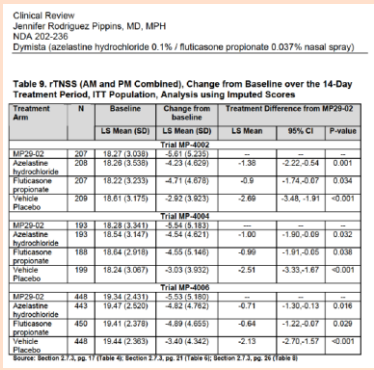
Journal Articles



DRUG EVALUATION

Intranasal Fluticasone Propionate
A Review of its Pharmacodynamic and Pharmacokinetic Properties, and Therapeutic Potential in Allergic Rhinitis

Harriet M. Bryson and Diana Faulds
Adis International Limited, Auckland, New Zealand



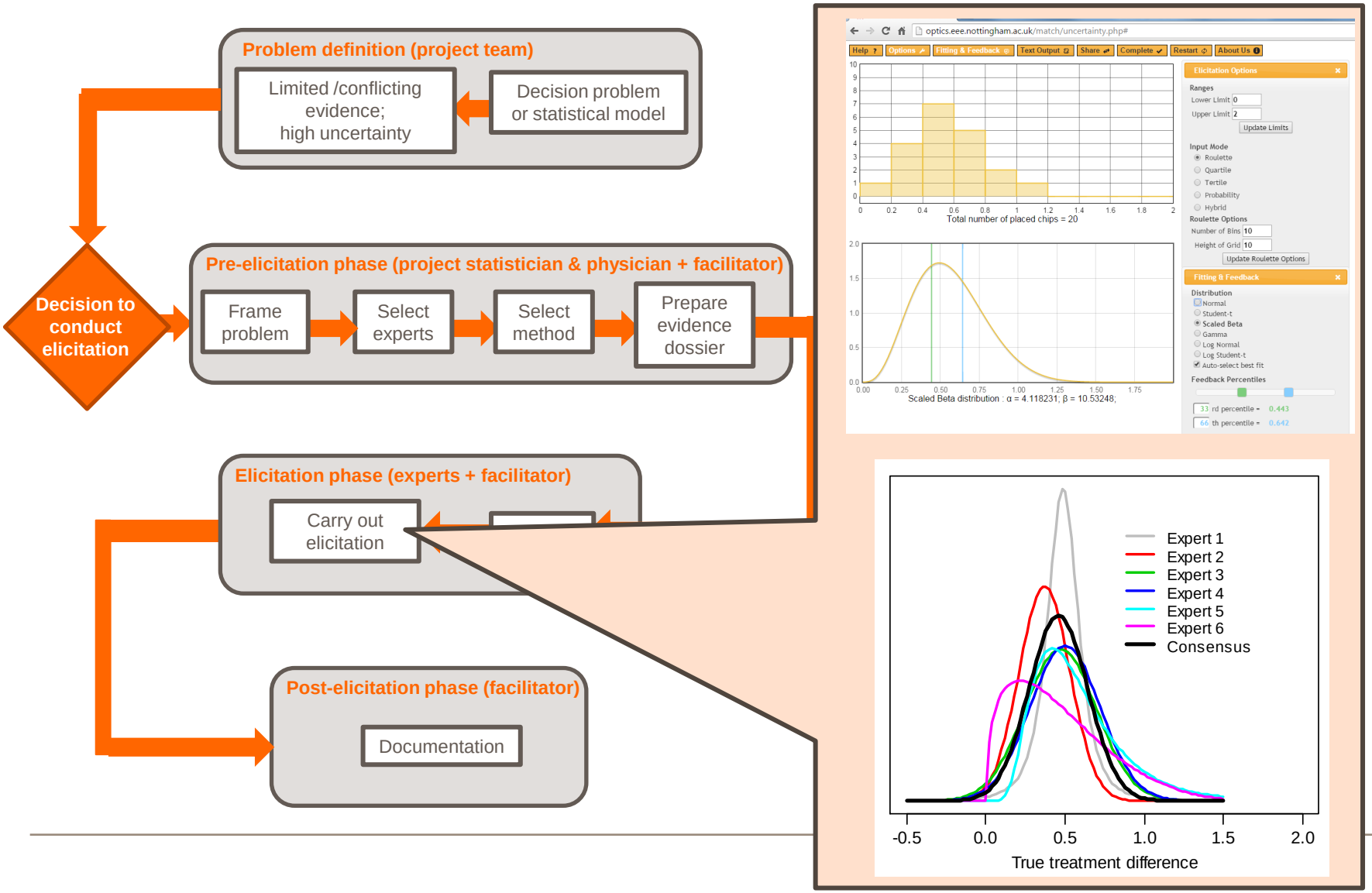
Clinical Review
Jennifer Rodriguez Pippins, MD, MPH
NDA 202-236
Dymista (azelastine hydrochloride 0.1% / fluticasone propionate 0.037% nasal spray)

Table 9. rTNSS (AM and PM Combined), Change from Baseline over the 14-Day Treatment Period, ITT Population, Analysis using Imputed Scores

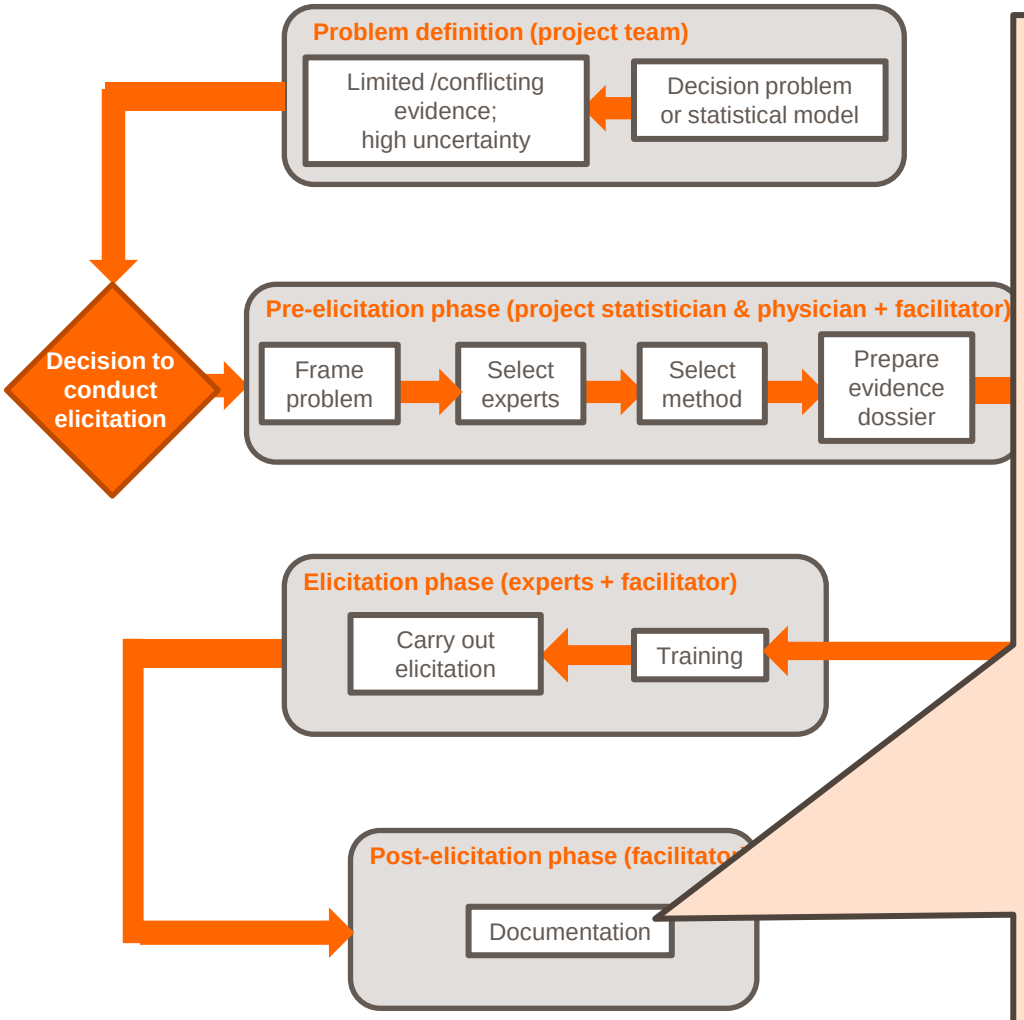
Treatment Arm	N	Baseline	Change from baseline	Treatment Difference from MP29-02		
		LS Mean (SD)	LS Mean (SD)	LS Mean	95% CI	P-value
Trial MP-4002						
MP29-02	207	18.27 (3.038)	-5.63 (5.239)	--	--	--
Azelastine hydrochloride	208	18.26 (3.538)	-4.23 (4.629)	-1.38	-2.22,-0.54	0.001
Fluticasone propionate	207	18.22 (3.233)	-4.71 (4.678)	-0.9	-1.74,-0.07	0.034
Vehicle Placebo	209	18.61 (3.175)	-2.92 (3.923)	-2.69	-3.48, -1.91	<0.001
Trial MP-4004						
MP29-02	193	18.28 (3.341)	-5.54 (5.193)	--	--	--
Azelastine hydrochloride	193	18.54 (3.147)	-4.54 (4.621)	-1.00	-1.90,-0.09	0.032
Fluticasone propionate	188	18.64 (2.918)	-4.55 (5.146)	-0.99	-1.91,-0.05	0.038
Vehicle Placebo	199	18.24 (3.067)	-3.03 (3.932)	-2.51	-3.33,-1.67	<0.001
Trial MP-4006						
MP29-02	448	19.34 (2.431)	-5.53 (5.180)	--	--	--
Azelastine hydrochloride	443	19.47 (2.520)	-4.82 (4.762)	-0.71	-1.30,-0.13	0.016
Fluticasone propionate	450	19.41 (2.378)	-4.89 (4.655)	-0.64	-1.22,-0.07	0.029
Vehicle Placebo	448	19.44 (2.383)	-3.40 (4.342)	-2.13	-2.70,-1.57	<0.001

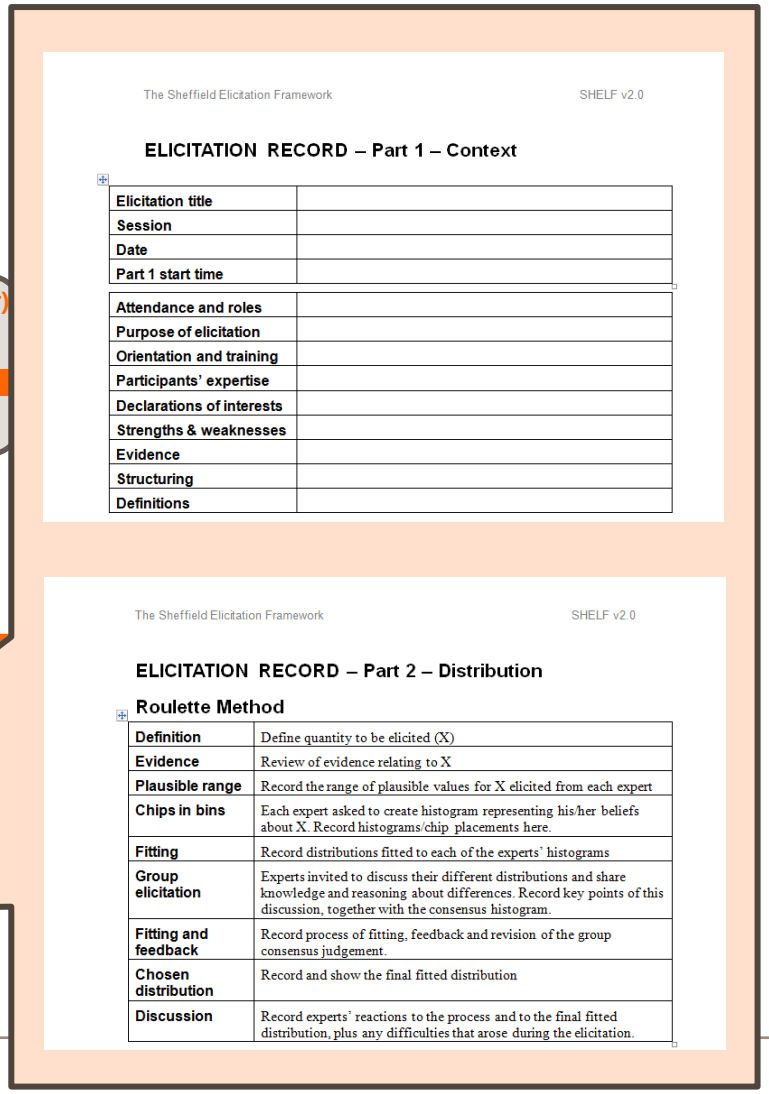
Source: Section 2.7.3, pg. 17 (Table 4); Section 2.7.3, pg. 21 (Table 5); Section 2.7.3, pg. 25 (Table 6)

Example of Prior Elicitation at GSK



Example of Prior Elicitation at GSK





The Sheffield Elicitation Framework SHELF v2.0

ELICITATION RECORD – Part 1 – Context

Elicitation title	
Session	
Date	
Part 1 start time	
Attendance and roles	
Purpose of elicitation	
Orientation and training	
Participants' expertise	
Declarations of interests	
Strengths & weaknesses	
Evidence	
Structuring	
Definitions	

The Sheffield Elicitation Framework SHELF v2.0

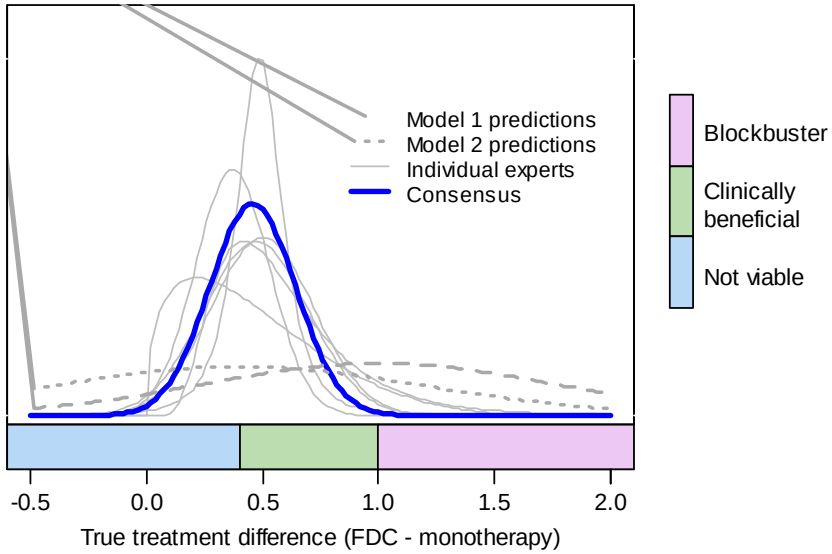
ELICITATION RECORD – Part 2 – Distribution

Roulette Method

Definition	Define quantity to be elicited (X)
Evidence	Review of evidence relating to X
Plausible range	Record the range of plausible values for X elicited from each expert
Chips in bins	Each expert asked to create histogram representing his/her beliefs about X. Record histograms/chip placements here.
Fitting	Record distributions fitted to each of the experts' histograms
Group elicitation	Experts invited to discuss their different distributions and share knowledge and reasoning about differences. Record key points of this discussion, together with the consensus histogram.
Fitting and feedback	Record process of fitting, feedback and revision of the group consensus judgement.
Chosen distribution	Record and show the final fitted distribution
Discussion	Record experts' reactions to the process and to the final fitted distribution, plus any difficulties that arose during the elicitation.

Communicating priors to decision makers

Belief distribution about true size of treatment effect

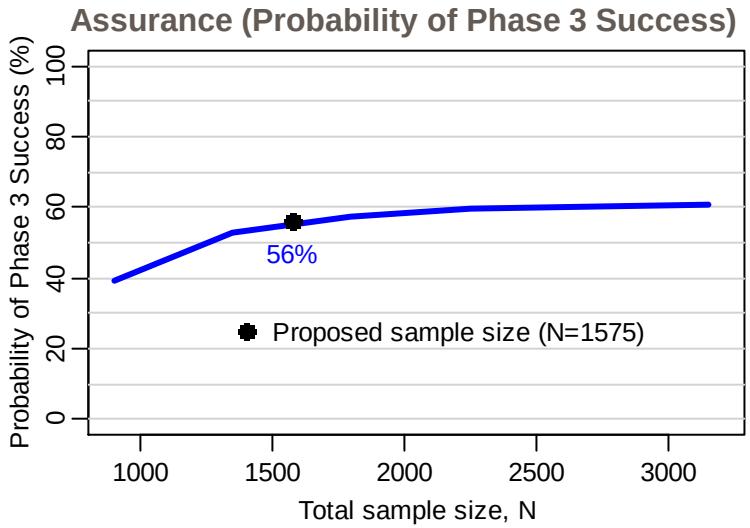


❑ **Model-based predictions**

- Multiple uncertainties in statistical model
- Available data insufficient to estimate parameters well
- Low precision for predicting phase 3 treatment effect

❑ Consensus belief distribution

- More informative than model-based prior, based on experts' knowledge in addition to available data
- Strong conviction that FDC could not lead to true outcome being worse than monotherapy
- Treatment effects > 1 would be exceptional



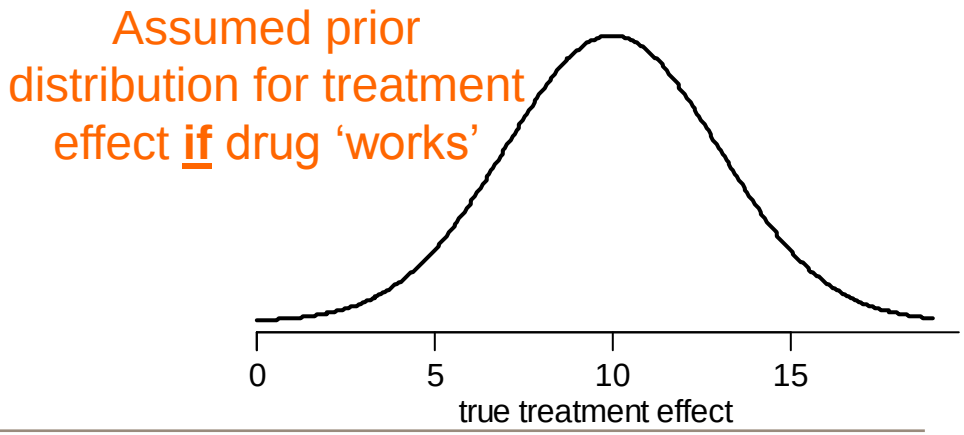
Success = p < 0.05 and observed effect > 0.4 in both P3 trials

❑ Sample sizes above ~1500 per arm yield negligible gains in assurance

❑ Plot shows assurance for 3:3:1:1 randomisation ratio; alternative designs with different randomisation ratios gave almost identical assurance values

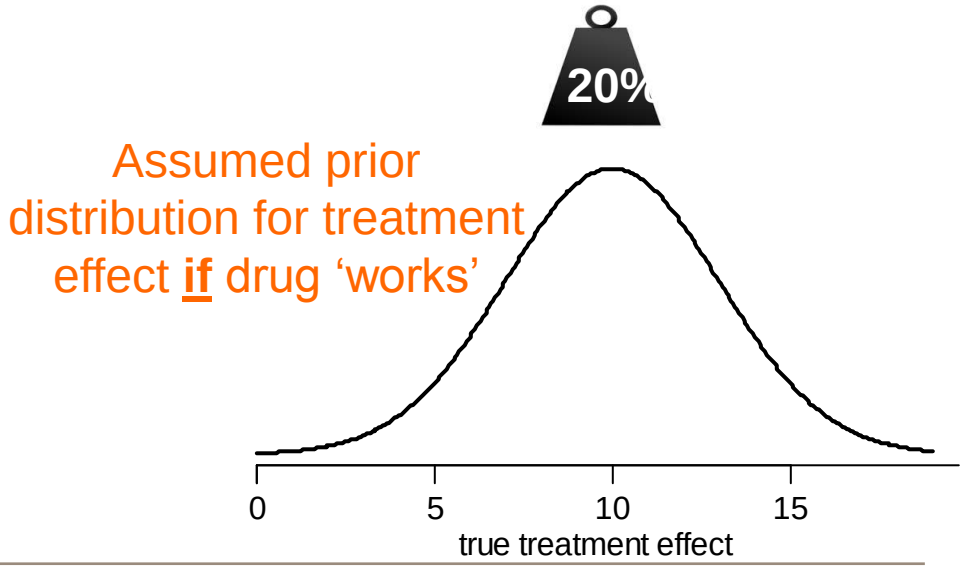
Managing the tendency for over-optimism in expert opinion

1. Elicit a prior for the true treatment effect conditional on the drug 'working' (e.g. mechanism translating)



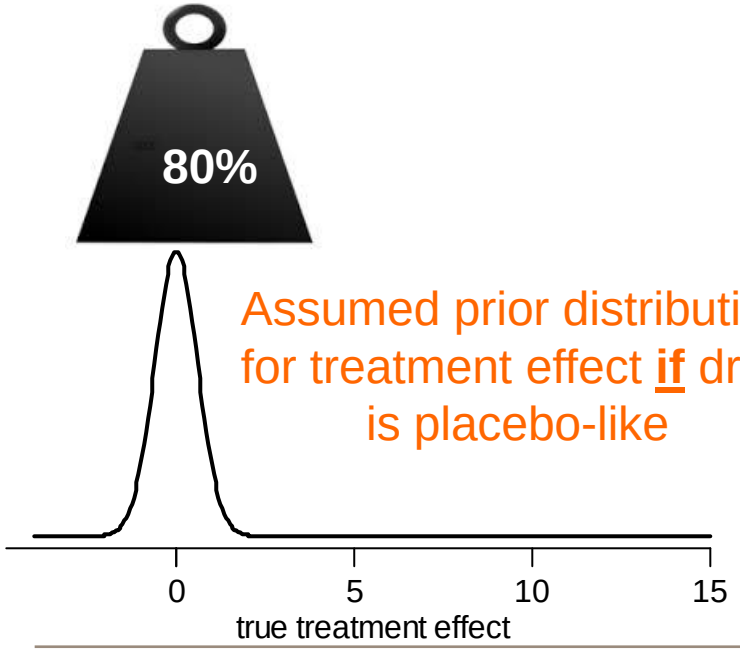
Managing the tendency for over-optimism in expert opinion

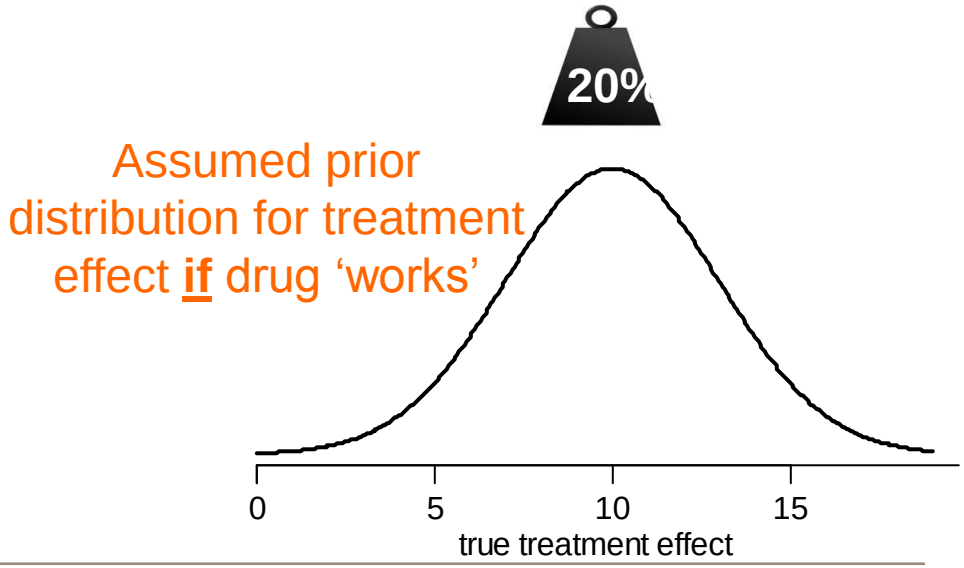
1. Elicit a prior for the true treatment effect conditional on the drug 'working' (e.g. mechanism translating)
2. Elicit a prior probability that the drug 'works'



Managing the tendency for over-optimism in expert opinion

1. Elicit a prior for the true treatment effect conditional on the drug 'working' (e.g. mechanism translating)
2. Elicit a prior probability that the drug 'works'
3. Combine with 'placebo-like' distribution tightly centred around zero

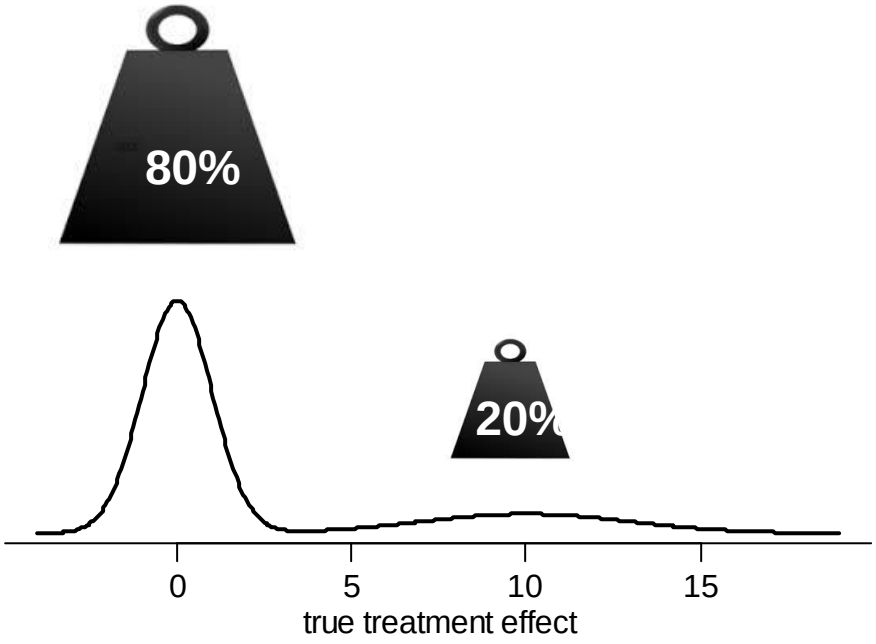




Managing the tendency for over-optimism in expert opinion

1. Elicit a prior for the true treatment effect conditional on the drug 'working' (e.g. mechanism translating)
2. Elicit a prior probability that the drug 'works'
3. Combine with 'placebo-like' distribution tightly centred around zero

→ Mixture prior



Example of Bimodal Prior Elicitation

Problem definition

Decision problem:

- Rare disease with history of studies failing in this disease area
- Ongoing Phase 2 study
- Early stages of planning Phase 3

Elicitation Aim:

- Elicit experts beliefs without the ‘bias’ of observing the phase II study
- Combine the prior with the observed phase II data so as to calculate the assurance for potential phase III designs

Example of Bimodal Prior Elicitation

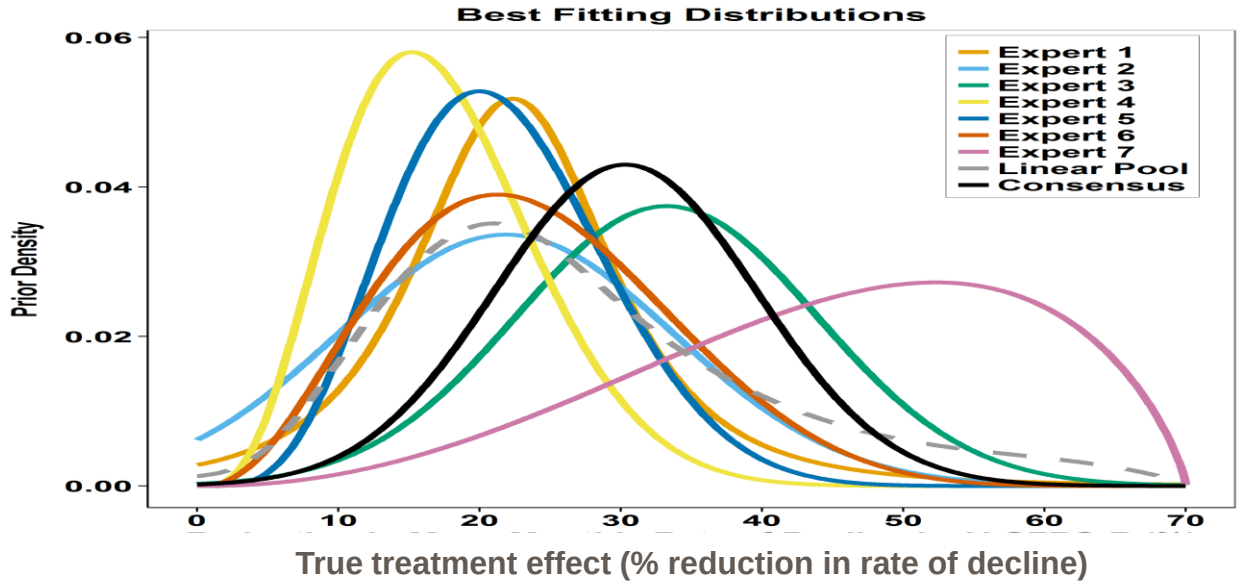
Elicitation

1. Prior belief that drug works (‘causes some relevant biological activity’)
   – Consensus was 25% (range: 10 to 40%)

Example of Bimodal Prior Elicitation

Elicitation

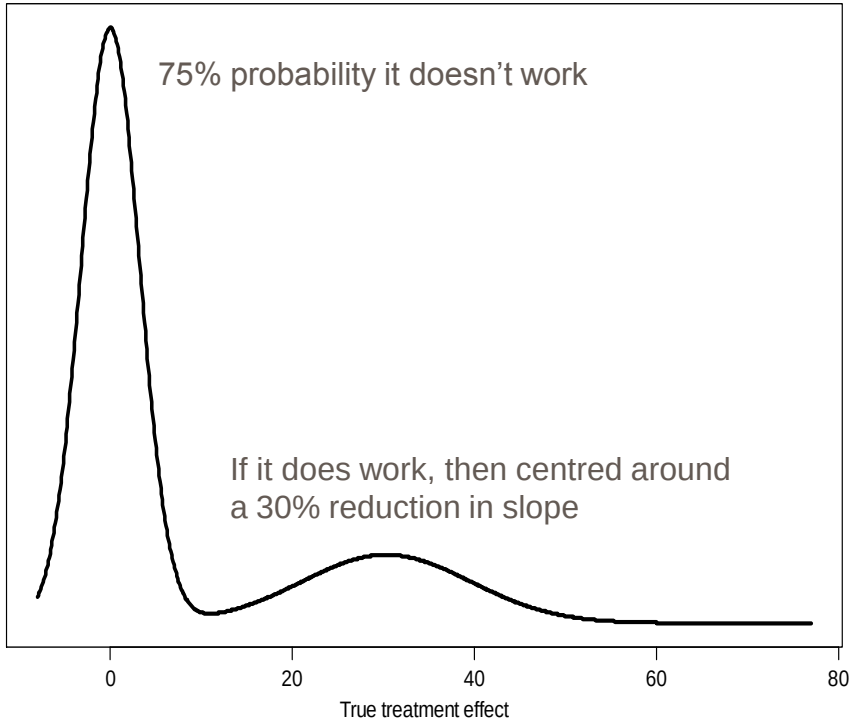
1. Prior belief that drug works ('causes some relevant biological activity')
   - Consensus was 25% (range: 10 to 40%)
2. Conditional on drug working, how efficacious is it?



Example of Bimodal Prior Elicitation

Overall mixture prior

- Update this with phase 2 data
- Can make statements about the posterior of the phase 2
- Use in assurance calculations for planning phase 3



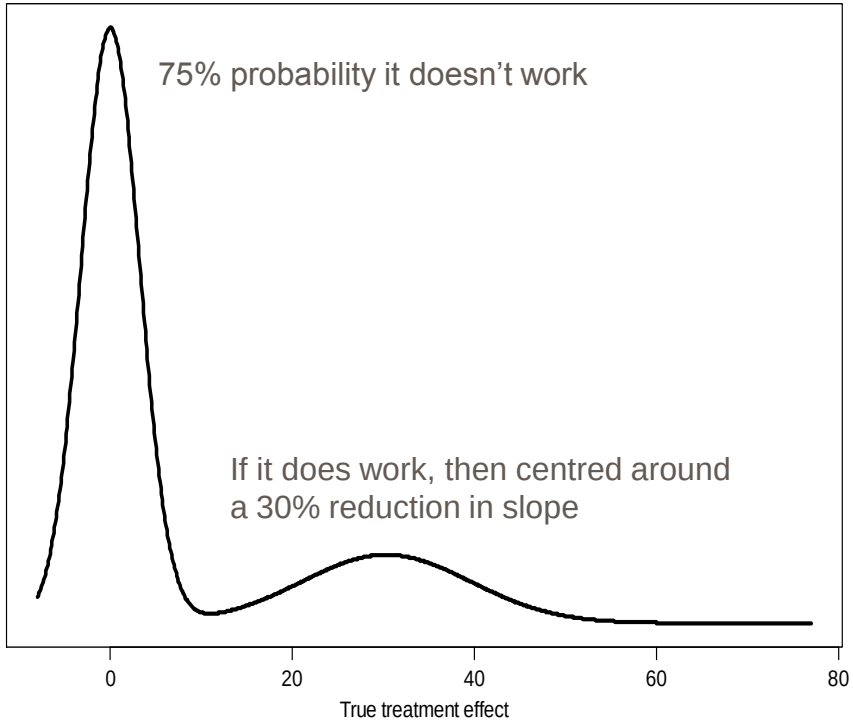
Example of Bimodal Prior Elicitation

Overall mixture prior

- Update this with phase 2 data
- Can make statements about the posterior of the phase 2
- Use in assurance calculations for planning phase 3

What actually happened....

- Phase 2 results were negative
  - ➢ Planning for Phase 3 did not go ahead
- Retrospective assurance calculation for Phase 2 study: **assurance=21%**
  - ➢ Should we have planned interim futility analysis?



Challenges and Benefits of Prior Elicitation

- Prior elicitation enables project teams to utilize historical data, prior knowledge from experts, and collective thought for a more robust output on study design and/or analysis
- 13 elicitations conducted at GSK to date
  - positive feedback received from all teams

Challenges and Benefits of Prior Elicitation

- Prior elicitation enables project teams to utilize historical data, prior knowledge from experts, and collective thought for a more robust output on study design and/or analysis
- 13 elicitations conducted at GSK to date
  - positive feedback received from all teams

Practical challenges:

- Experienced, skilled facilitators are essential
- Need at least 2 facilitators, one to lead and one to run software and keep written record of elicitation session
- Logistics extremely challenging
  - 3-6 hour time commitment
  - Face-to-face in same room (VTC an option but not ideal)
- Training of experts is essential
- Need experts who are open-minded

Challenges and Benefits of Prior Elicitation

Benefits:

- Assurances of key outcomes are what decision makers need
  - Power is more or less useless for decision making
  - But you have to bite the bullet of characterising knowledge and uncertainty about true effects → prior distributions

Challenges and Benefits of Prior Elicitation

Benefits:

- Assurances of key outcomes are what decision makers need
  - Power is more or less useless for decision making
  - But you have to bite the bullet of characterising knowledge and uncertainty about true effects → prior distributions
- Formal and rigorous elicitation of expert knowledge enables
  - honest reflection of current state of knowledge
  - robust basis for conversations about asset management, trial design and interpretation of trial findings

Challenges and Benefits of Prior Elicitation

Benefits:

- Assurances of key outcomes are what decision makers need
  - Power is more or less useless for decision making
  - But you have to bite the bullet of characterising knowledge and uncertainty about true effects → prior distributions
- Formal and rigorous elicitation of expert knowledge enables
  - honest reflection of current state of knowledge
  - robust basis for conversations about asset management, trial design and interpretation of trial findings
- All teams at GSK are being encouraged to explore the potential of Prior Elicitation for their projects
  - "One pager" summarising Prior distribution + Assurance required for all major governance board milestones

Challenges and Benefits of Prior Elicitation

Benefits:

- Assurances of key outcomes are what decision makers need
  - Power is more or less useless for decision making
  - But you have to bite the bullet of characterising knowledge and uncertainty about true effects → prior distributions
- Formal and rigorous elicitation of expert knowledge enables
  - honest reflection of current state of knowledge
  - robust basis for conversations about asset management, trial design and interpretation of trial findings
- All teams at GSK are being encouraged to explore the potential of Prior Elicitation for their projects
  - "One pager" summarising Prior distribution + Assurance required for all major governance board milestones
- Impact
  - 25% reduction in a P3 study size (saving >£15M and 8 months)
  - Inclusion of interim futility analyses in several studies

Acknowledgements

Nigel Dallow
Tim Montague
Adam Crisp
John Davies
Graeme Archer
Sam Miller
Maria Costa
Will Powley
Jane Temple
Jon Robertson
Kim Hacquoil
Nicola Scott
Frank Fang
Younan Chen
Grace Zhang
Faiz Ahmad

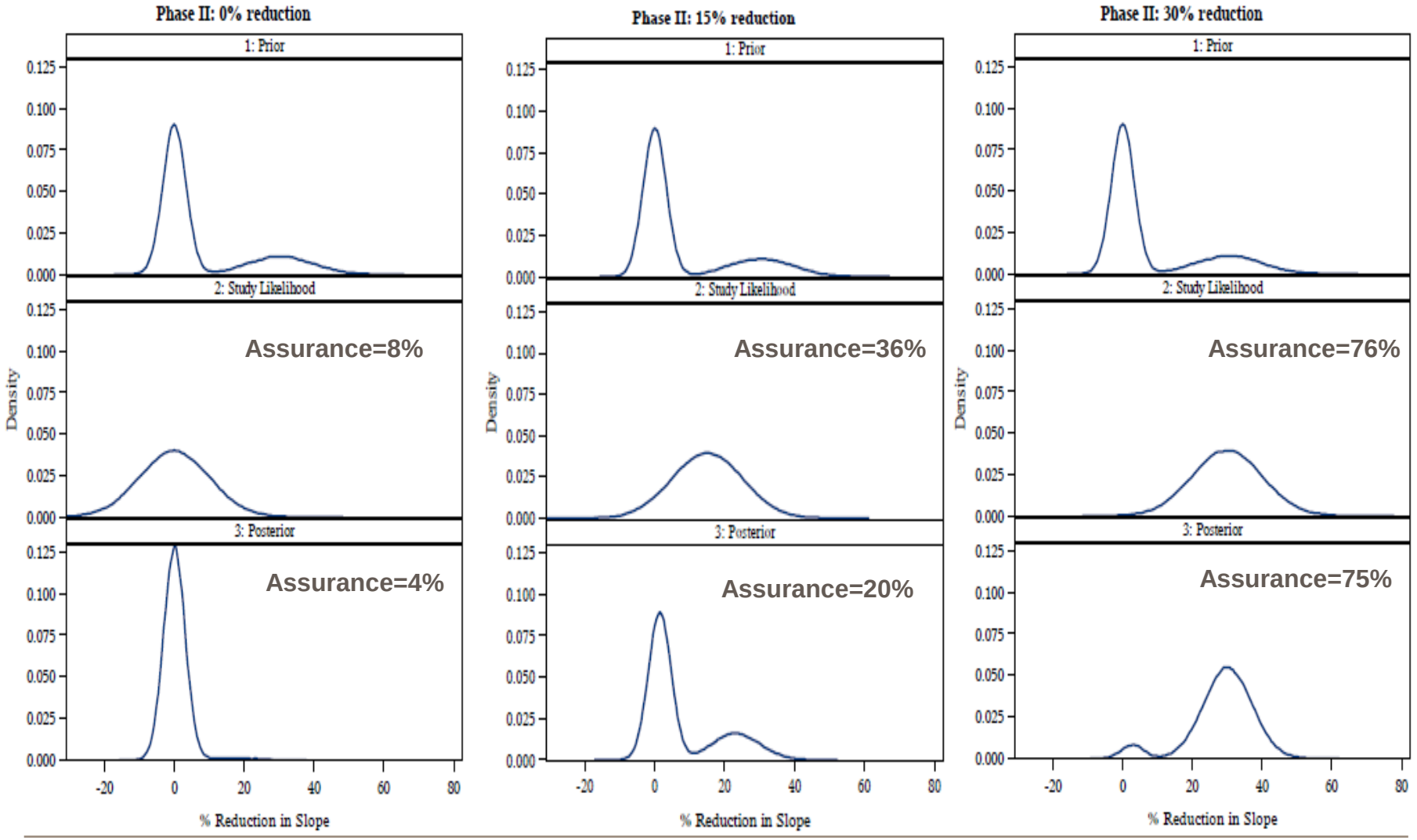
Sara Hughes

All of GSK Clinical Statistics

Thank you for listening

Backups

Assurance for Phase 3 Design – Possible Scenarios



Assurance for Phase 3 Design